CROWDSOURCED INCIDENT AND CONGESTION DATA

Description
- Travelers report incidents and congestion in real time using smartphone apps or similar technology.

Applications
- Managing traffic in real time through awareness of new incidents and congestion.
- Can be archived and used for planning and performance management.

Attributes
- Low latency and less detailed.
- More prevalent and accurate in dense urban areas, highly traveled corridors, and regions with a high level of technology acceptance.

Data Availability
- Mining/monitoring social network data.
- Citizens posting maintenance issues for agency action.

Pros and Cons
- Tends to provide broader coverage.
- Can be timelier than agency-generated data.

Source: Tweet Deck screenshot provided courtesy of the MATOC Program.
Crowdsourced Data Applications at the Port Authority of New York and New Jersey
ROADSIDE BASIC SAFETY MESSAGE (BSM) DATA

Description
- Packets of data with vehicle position, heading, speed, and other information relating to a vehicle’s state and predicted path.
- Some of the modern roadside equipment can collect BSM sent from equipped vehicles.

Applications
- Vehicle platooning.
- Queue warning.
- Intelligent traffic signal systems and transit signal priority.

Data Availability
- Pilot and test sites – United States Department of Transportation (USDOT).
- A neutral third party uses APIs to collect and transmit data.

Pros and Cons
- Large data sets are challenging to process and use in real time.
- Unknown data aggregation levels needed in various applications.
- Low market penetration and minimal use for real-time operations.

Source: U.S. DOT, Intelligent Transportation Systems Joint Program Office.
REAL-TIME AND ARCHIVED TRAJECTORY DATA

Description
- Time-stamped location data from moving vehicles, cell phones, or other GPS-enabled devices.

Application
- Real-time traffic pattern analysis.
- Multi-modal system utilization.
- Operations planning.

Attributes
- Unique device, vehicle, and trip IDs.
- Trip departure (origin) and arrival (destination) times and locations.
- Periodic trip “waypoints.”

Data Availability
- Relatively new to the market with few companies providing this information.

Pros and Cons
- Can construct an entire trip based on origin, destination, and route.
- Lag time of 1 to 2 days between data collection and availability.
- Large data sets.

Original map: © 2017 Google.
PROBE-BASED SPEED DATA

Description
- Speed and travel time information from GPS devices or toll tags.

Application
- Monitoring real-time congestion, performance management, traveler information.

Attributes
- Latency ranging from 10 to 30 seconds, sufficient for most operations and planning purposes.
- Collection of speed and travel time for roadway segments.
- Continuous improvement of quality and coverage.

Data Availability
- Toll facility tag readers can collect probe speed data.

Pros and Cons
- Ubiquitous roadway network coverage and data collection for equipped vehicles or devices.
- High quality data in metropolitan areas, but lower quality data on arterial roadways due to interrupted flow.
- Lack of associated volume information.

Source: University of Maryland CATT Lab.
**WI-FI AND BLUETOOTH RE-IDENTIFICATION DATA**

**Description**
- Location data collected by static, distributed Wi-Fi and Bluetooth scanners used to infer speed and travel time information; origin-destination data possible with dense network of sensors.

**Application**
- Origin-destination analysis.
- Traffic signal optimization.
- Work zone management.
- Real-time queue warnings.

**Attributes**
- Data consists of a unique media access control (MAC) address for identification.

**Data Availability**
- Several private companies provide this data.

**Pros and Cons**
- Provides insight into individual trips between sensors.
- Sensors are inexpensive and lightweight but require power and communications and are costly to deploy widely.
- Built-in data filtering can remove vehicle data when off the network for extended times.
- Wi-Fi pollution and general interference can result in distortion or data drops.
CREDIT CARD TRANSACTIONS DATA

Description
- Credit card swipe data that serves as a surrogate measure of power outage information.

Application
- Locate open businesses that provide fuel, food, medications, and medical supplies, during a prolonged power outage.

Attributes
- The data only has the number of transactions and time of each transaction, but no personally identifiable information.
- Positioning information (addresses and business names) ties transactions to location.

Data Availability
- Service is available 24/7 and near real-time data delivery.

Pros and Cons
- Only useful during major regional power events.
- Data providers require additional security levels and credentialing for users.

CONNECTED VEHICLE DATA FROM THIRD PARTIES

Description
- Data from vehicles including wiper use, headlight use, heavy braking, traction control, fuel consumption, emissions, travel speed, acceleration and deceleration, etc.

Application
- Rain and micro-level weather predictions.
- End-of-queue or debris detection.
- Slippery or icy road condition detection.

Data Availability
- This may soon be available to DOTs as part of their regular incident data feeds.

Pros and Cons
- Enhanced situational awareness.
- Allows for pre-positioning of maintenance and response crews to mitigate impact.
- Can report potentially dangerous events and surface conditions.
- Data could be voluminous and distracting to operators.

Source: University of Maryland CATT Lab.
HIGH RESOLUTION SIGNAL DATA

Description
- Includes more than 150 different traffic signal controller events recorded at a very high frequency.

Application
- Real-time signal system performance measurement to automatically adapt to changing conditions.
- Detection and recording of red light and right-turn-on-red violations.
- Archived data can provide insights for planning.
- Supports performance measures – percentage of arrivals on green, capacity utilization at intersections, red-light and back-of-queue collision reduction.

Pros and Cons
- Provides data at high frequency, allowing prediction of congestion and signal performance but also requires appropriate expertise and infrastructure to store and process.
- Signal controllers are inexpensive and can be phased into existing systems, allowing quick and affordable deployment.
IN THE SPOTLIGHT...

Use of Signal Data

Mark Taylor

Utah Department of Transportation
ROADWAY WEATHER PREDICTIONS

Description
- Ground-based (i.e., at street level) near-term weather predictions updated hourly using a mix of satellites, ground-based radar, ground sensors, and connected vehicle sensors.

Application
- Helps to optimize and improve weather response, reduce staffing, fuel and chemical costs, and pinpoint treatment applications.
- Improving safety on roads by indicating dangerous weather events.

Attributes
- Precipitation type, rate, and accumulations.
- Air temperature and wind speed and direction.

Data Availability
- Available from the private sector.
- Some in-house, full-time agency meteorologists.

Pros and Cons
- Beneficial where hazardous weather can affect mobility and safety.
- Reduces maintenance costs, excess materials, fuel consumption, and optimizes resource management.
- A high cost for quality road weather data, potentially limited geographical resolution, and operator impacts.
COMPUTER AIDED DISPATCH (CAD) DATA

Description
- Emergency location and type information.

Application
- Helps first responders and safety agencies identify emergency type and location.
- Quick identification of incidents, including where a TMC lacks coverage or jurisdiction.

Attributes
- Information is real-time and is pushed out to responders’ mobile data terminals.
- Includes incident type, timestamp, and dispatched unit information.
- Some data also contain location, actions taken, caller information, and status.

Data Availability
- Agencies work with public safety partners to exchange data, since it is not available everywhere.

Pros and Cons
- Faster incident detection and more effective response and equipment deployment.
- Automated exchange reduces communication workload and increases situational awareness.
- Data collected for all calls, not just transportation calls, so feeds must be filtered.
- Often uses abbreviations, shorthand, and free text requiring interpretation.
- Free text entries can include errors that invalidate data so it is ignored or discarded.
CURRENT AND EMERGING PRIVATE-SECTOR BUSINESS MODELS

BUSINESS MODELS FOR SELLING DATA AND SERVICES

Data providers have many different business models for selling data and services including revenue sharing with agencies and exclusive access to data.

Pricing structure is based on:

- Centerline miles.
- Population of the covered region.
- Number of requests to application programming interfaces (APIs).
- Number of users.
- Number of “studies” conducted with the data.
- Bundling of data or services (discounts).
VALUE OF AGENCY DATA

- Most private sector data providers still want agency-produced datasets.
- The type of data deemed most valuable has changed.
- Previously, traveler information providers and the media valued volume and speed sensor data. With the growing penetration of probes, this data is now less attractive.
- Computer aided dispatch data is prized for the speed of data retrieval and improved coverage.
AGENCY DATA VALUATION

- Identify the opportunity.
- Assess your uniqueness.
- Assess the market.
- Tiers of service.
- Evaluate your costs.
- Profit or cost-recovery.
- Test the market.